

# Damning probes find Instagram is key link connecting pedophile rings

## reader comments

51 with

Instagram has emerged as the most important platform for buyers and sellers of underage sex content, according to investigations from The Wall Street Journal, Stanford Internet Observatory, and the University of Massachusetts Amherst (UMass) Rescue Lab.

While other platforms play a role in processing payments and delivering content, Instagram is where hundreds of thousands—and perhaps millions—of users search explicit hashtags to uncover illegal “menus” of content that can then be commissioned. Content on offer includes disturbing imagery of children self-harming, “incest toddlers,” and minors performing sex acts with animals, as well as opportunities for buyers to arrange illicit meetups with children, the Journal reported.

Because the child sexual abuse material (CSAM) itself is not hosted on Instagram, platform-owner Meta has a harder time detecting and removing these users. Researchers found that even when Meta’s trust and safety team does ban users, their efforts are “directly undercut” by Instagram’s recommendation system—which allows the networks to quickly reassemble under “backup” accounts that are usually listed in the bios of original accounts for just that purpose of surviving bans.

A Meta spokesperson told Ars that the company works “aggressively to fight” child exploitation on all its platforms “and to support law enforcement in its efforts to arrest and prosecute the criminals behind it.” Because criminals’ tactics “constantly change,” Meta said it has enacted “strict policies and technology to prevent them from finding or interacting with teens on our apps” and hired “specialist teams who focus on understanding their evolving behaviors so we can eliminate abusive networks.” These efforts led Meta to dismantle 27 abusive networks between 2020 and 2022 and ban 490,000 accounts violating child safety policies in January 2023, the spokesperson reported.

But these tactics do not appear to be doing enough to combat the problem, researchers said. UMass Rescue Lab director Brian Levine told Ars that it took his team minutes to uncover pedophile rings operating on Instagram after identifying “simple tags” used to help connect buyers and sellers. The Wall Street Journal reported that the hashtags researchers identified could be obvious, like “pedowhore,” or rely on code words, like “cheese pizza,” which shares initials to allude to child pornography.

#### Advertisement

Levine said that since Instagram’s trust and safety team has broader access to search the platform, it should be able to monitor for hashtags more effectively than outside researchers. However, the team does not seem to be effectively monitoring the situation, as it’s missing easily discoverable hashtags.

This was the first time Levine’s lab looked into Instagram as part of the team’s work to research online child victimization and build tools to prevent it, Levine told Ars.

“I think their trust and safety team needs a lot of help,” Levine added.

In research published yesterday investigating self-generated CSAM (SG-CSAM) allegedly posted by minors advertising their own content, Alex Stamos, Stanford Internet Observatory (SI0) director, told Ars that his team didn't specifically target Instagram. But the team found that the platform's recommendation system "plays a particularly important role as the discovery mechanism that introduces buyers to sellers." Levine told Ars that Instagram has an obligation to ensure its recommendation system does not promote abusive content.

The Journal revealed that Instagram repeatedly failed to ban hashtags, take down content, and prevent promotion once content was detected. After the Journal sent inquiries, for example, Meta confirmed that it was "in the process of removing" hashtags promoting illegal materials like "pedobait" or "mnsfw" (minor not safe for work) and admitted that it "permitted users to search for terms that its own algorithms know may be associated with illegal material." A Meta spokesperson told Ars that it has already "restricted thousands of additional search terms and hashtags on Instagram."

The WSJ reported that Instagram was so ineffective at stopping pedophile rings from forming that it would sometimes display a pop-up screen when users searched for known abusive content, warning users that "these results may contain images of child sexual abuse." Users could then choose to "see results anyway" or "get resources" regarding the harms of consumption of abusive content. Meta told the Journal that it was removing this option but failed to explain why the option was ever given in the first place.

Meta also admitted that it has repeatedly failed to act on reports of child abuse on Instagram because of "a software glitch" preventing reports from ever being processed. This bug has reportedly since been fixed. Now Meta said it will provide additional training for content reviewers and assemble an internal task force "to investigate these claims and

immediately address them.”

<https://arstechnica.com/?p=1946320>